**Bioinformatics Exercise 6:** GeneCards database for finding information about a gene

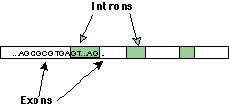
In agreement with evolutionary principles, scientific research to date has shown that all genes share common elements. For many genetic elements, it has been possible to construct consensus sequences, those sequences best representing the norm for a given class of organisms (e.g, bacteria, eukaroytes). Common genetic elements include promoters, enhancers, polyadenylation signal sequences and protein binding sites.

A mRNA can be divided into three parts: a 5′ untranslated region (5′ UTR), the polypeptide coding region, sometimes called the open reading frame (ORF), and the 3′ translated region (3′ UTR).

The first codon in a messenger RNA sequence is almost always AUG. While this reduces the number of candidate codons, the reading frame of the sequence must also be taken into consideration.

There are six reading frames possible for a given DNA sequence, three on each strand, that must be considered, unless further information is available. Since genes are transcribed away from their promoters, the definitive location of this element can reduce the number of possible frames to three. The location of the appropriate start codon will include a frame in which there are no apparent abrupt stop codons. Incorrect reading frames usually predict relatively short peptide sequences. Therefore, it might seem deceptively simple to ascertain the correct frame. In bacteria, such is frequently the case. However, eukaryotes add a new obstacle to this process, introns.

Intron/exon splice sites can be predicted on the basis of their common features. Most introns begin with the nucleotides GT and end with the nucleotides AG.



1. Have a look at the organization of introns/exons on the [DMD homepage](http://www.dmd.nl/seqs/murefDMD.html).
2. GeneCards is a database of human genes, their products and their involvement in diseases and there are several ways to search:

* Go to the [GeneCards website](http://www.genecards.org/) .
* Type DMD into the search box.
* Click on “Disorders” under the search box.

This card shows all information available on the gene, with links to corresponding databases, e.g.

1. Find the link to OMIM this gene on the GeneCard page.
2. **What can you find out about this disorder at these sites?**